IN THE SPECIFICATION:

Please replace paragraph [0110] with the following paragraph:

The invention employs two-dimensional models with forms like $\log(Y_{ij}) | i \in S_m, j \in G_l \sim N\left[f\left(t_{il}, \alpha_{mi}, \beta_{lj}, \gamma_{ml} \right), \sigma^2 \right]$, where i and j index the expression data by gene and sample respectively, m and l index latent classes on the corresponding dimensions, the t refer to gene-specific intensity parameters, and various forms for the function f are chosen. The following analytic results are based on an additive form for f, α_{mi}, β_{lj}, γ_{ml}, that is, an ANOVA-like additive model is used for incorporating main effects and an interaction term for the latent class means. Using its complete conditional distributions in the presence of weak prior information, a modified Metropolis algorithm is employed to converge to draws from the posterior distribution of this model. No information about the known cancer classification of the tissues is provided to this model, since the objective is to illustrate the SPAM model's ability to differentiate and quantify molecular fingerprints of differing tissue types.

Please delete paragraph [0028] from the specification.

Please replace paragraph [0029] with the following paragraph:

[0029] Fig. 1 displays comparative results from TaqMan and Microarray assays of the fibroblast data for 5 expressed genes.

Please replace paragraph [0030] with the following paragraph:

[0030] Fig. 2 shows adjusted estimates of Mast and B4-2 expression derived by the present invention from the microarray data.

Please delete paragraph [0031] from the specification.

Please replace paragraph [0032] with the following paragraph:

[0032] Fig. 3 displays hierarchical clustering of time-course data for ischemic decay. (A) shows clustering results using the average of the three replicas for each time point. The dendogram shows the time points related by progression. The 40 and 60 minute time points are most closely related to each other, and least related to the earliest time points. Similar results emerge when the measurements at each time point for each gene are compared to the measured valued for that gene at (B) the 5 minute of (C) the 60 minute measurement. These results suggest a change in apparent expression level over time with an increasing deviation from the in-vivo measurement at greater ischemic times.

Please delete paragraph [0033] from the specification.

Please replace paragraphs [0034] – [0043] with the following paragraphs:

[0034] Fig. 4 illustrates a CLONTECH filter.

[0035] Fig. 5 illustrates a portion of a NEN MICROMAX slide.

[0036] Fig. 6 illustrates a portion of an AFFYMETRIX chip.

[0037] Fig. 7 shows the estimated assignment of each of a set of tissues into each of a set of latest tissue classes as estimated by the present invention.

[0038] Fig. 8 shows in interaction estimates (γ_{ml}) between latent gene classes and latent tissue classes, as estimated by the present invention.

[0039] Fig. 9 shows a representative 2-D gel of proteins after exposure to peroxisome proliferators.

[0040] Fig. 10 shows phosphorylated and non-phosphorylated versions of a protein in a 2-D gel.

[0041] Fig. 11 shows a flow diagram of a computer implemented process of the present invention.

[0042] Fig. 12 illustrates a process for Bayesian estimation employed by the present invention.

[0043] Fig. 13 illustrates multi-chain monitored algorithms employed by the present invention.

Please delete paragraph [0074] from the specification.

Please delete paragraph [0075] from the specification.

Please replace paragraph [0077] with the following paragraph:

[0077] The invention makes it possible to adjust for known issues with microarray estimation of true expression relative to Northern blot experiments. By way of illustration the published TaqMan assay and microarray results of the above fibroflast data for expressed genes are reproduced in Fig. 1. Although there is substantial correlation between the results, they are quite different. Specifically, the microarray data is overall more variable across the experiment, underestimates the large changes in RNA expression, and suffers from edge effects. All three problems are illustrated by the Mast and B4-2 genes. For example, the TaqMan assay provides that there are substantial differences in expression between Mast and B4-2 from 4 hours on, at which time Mast is an order of magnitude more repressed that B4-2. Because the present invention pools information across gene products and microarray hybridizations, it uses biological correlations across measurements to reduce biases in particular observations. The present invention also adjusts for biases inherent to specific biological techniques.

Please replace paragraph [0078] with the following paragraph:

[0078] Fig. 2 shows adjusted estimates of the present invention of Mast and B4-2 expression derived from the microarray data. This approach is easily adaptable to other situations such as for example analyzing changes in expression of fibroblast and epithelial cells in response to selenomethionine, a compound known to decrease clinical risk of certain cancers in high-risk groups. Data from Northern blot and microarray data can be analyzed using the invention to verify and adjust expression quantitation.

Please replace paragraph [0087] with the following paragraph:

[0087] For each of the approximately 2,400 arrayed genes, the natural logarithm of the ratio of Cy5/Cy3 background-subtracted intensities is determined and used for subsequent analysis. The prior art includes performing gene and time-point cluster analysis using the Cluster/Tree View hierarchical clustering package (see incorporated reference Eisen, M.B., et al., Cluster analysis

and display of genome-wide expression patterns. Proc Natl Acad Sci USA, 1998 .92(25): pp. 14863-8), which uses Pearson correlation coefficients as a measure of similarity and average-linkage clustering. Clustering results are displayed visually, with each experimental time-point represented by a column in the display and each gene by a separate row. In the prior art Eisen reference, elements of the display are colored to represent its mean-adjusted ratio value; red-colored and green-colored cells represent, respectively, higher and lower levels of expression relative to the test sample; relative expression is represented by the relative brightness of the signal.

Please replace paragraph [0088] with the following paragraph:

[0088] Of the 2,400 genes in the arrays at least 2,114 provide useful data in at least one experiment. Analysis of the 1,420 genes that gave signals above background for all eighteen measurements is performed. A time course, with genes clustered by temporal expression pattern can be shown. It can be shown that the temporal expression pattern of each gene is represented by a horizontal row; the expression patterns of all genes in a single experiment is represented by a column. Using this red/green display as proposed in the prior art, it is difficult to visually assess relative changes in gene expression levels over time.

Please delete paragraph [0089].

Please replace paragraph [0090] with the following paragraph:

[0090] Analysis by the prior art Eisen clustering analysis does identify some relationships suggesting that the earlier time points cluster away from their later time points (Fig. 3); however, unlike the present invention, the prior art method give no indication as to the reason for the clustering.

Please replace paragraph [0091] with the following paragraph:

[0091] Application of the time-course model using the present invention reveals that there are 3 patterns in the ischemia data. These three prevalent patterns account for 68.2%, 17.8% and 13.4% of the 1420 genes, respectively. Pattern I corresponds to an average of 27% over 60 minutes from 5 minute baseline level of expression. 63.8% of the genes with at least 80%

probability of membership in this pattern show average increases in expression over 60 minutes. The remainder decrease on average. Pattern II genes show the least ischemia-related effects, demonstrating an average of only 12% over 60 minutes. In contrast to pattern I, 67.5% of the genes with at least 80% probability of membership in this pattern are decreasing in expression on average over time. The remaining 32.5% in this pattern increase an average of 12% over 60 minutes. Finally, pattern III genes (13.4% of the sample) show the greatest sensitivity to ischemia, changing an average of 50% over 60 minutes, with about the same number increasing as are decreasing.

Please replace paragraph [0092] with the following paragraph:

[0092] In all these patterns, the null hypothesis of no change is not rejected with small sample size at any of the first 4 time points, from 5 through 20 minutes. In patterns I and III especially, 40 and 60 minute time points are found to deviate significantly from the 5 minute expressions.

Please replace paragraph [0094] with the following paragraph:

[0094] Based on these data, a number of clear conclusions are drawn. First, temporal changes in gene expression levels do occur following tissue excision, with detectable changes after as little as 20 minutes. This observation is in conflict with the conclusions one might draw regarding the "quality" of the RNA based on 18 s and 28 s bands seen on a gel. Typically, stability of RNA is usually determined by the integrity of the 18 s and 28 s ribosomal bands on an agarose gel. A gel showing a good 18 s:28 s ratio for tissue RNA samples obtained at all time points, would suggest that the RNA is stable.

Please replace paragraph [0107] with the following paragraph:

[0107] Microarray expression analysis is used to analyze differential gene expression patterns. For the general notion of gene expression analysis by using microarrays several well-known references exist in the art as enclosed herein by way of reference: The principles of such technologies are disclosed in U.S. Pat. Nos. 5,556,752, 5,744,305, 5,837,843, 5,843,655, 5,874,21 and 5,849,486. Published International (PCT) patent applications, such as WO 99/27137 and WO 99/10538, disclose additional information, all of which are incorporated by

reference herein. Illustrations of a CLON-TECH filter (Fig. 4), an NEN MICROMAX slide (FIG. 5) and an AFFYMETRIX chip (FIG. 6) are provided herewith.

Please replace paragraph [0112] with the following paragraph:

[0112] As a result, evidence is obtained to support the existence of up to 5 latent tissue classes in these data. Two of these classes contain relatively large proportions of normal or cancer tissues. The estimated assignment of each tissue into each latent tissue class in given in FIG. 7. Tissues are assigned based on having a greater than 80% estimated probability of membership in a particular class.

Please replace paragraph [0115] with the following paragraph:

[0115] The interaction between latent gene classes and latent tissue classes, as shown in FIG. 8, illustrates the ability of certain gene classes to discriminate among the estimated tissue classes. For example, this figure demonstrates that expressed genes that are primarily in latent gene class 7 are upregulated in most normal tissues relative to most cancer tissues in this data set; however, they are also highly upregulated in tissue classes III, IV and especially V, which have tumor as well as normal tissues. The latent tissue class III has many relative expression levels between those of tissue classes I and II, implying that it may consist of earlier stage cancer tissues and adjacent normal tissues with pre-malignant changes in gene expression.

Please replace paragraph [0125] with the following paragraph:

The LSG Kepler system involves an extensive two-dimensional mathematical filter that removes background, deconvolves each protein spot into one or more Gaussian peaks, and calculates the volumes under each peak (representing protein quantity). A multiple montage program allows the comparable areas of a series of up to 1,000 gels to be displayed and inter-compared visually to check on pattern matching. In matching individual gels to the chosen master 2-D pattern, a series of about 50 proteins is matched by an experienced operator working with a montage of all the 2-D patterns in the experiment. See FIG. 9 for a representative 2-D gel of proteins after exposure to peroxisome proliferators. As phosphorylated and un-phosphorylated versions of a protein occur at different locations on the 2-D gel, differential quantitation of the different forms is further assessed (FIG. 10). Subsequently, an automatic program is used to match additional

600-1000 spots to the master pattern using as a basis of the landmark data entered by the operator. A 2D-GE analysis of an individual tumor results in a protein molecular fingerprint which is directly compared to that of numerous other tumors.

Please replace paragraph [0149] with the following paragraph:

[0149] FIG. 11 provides a flow diagram for a computer implemented process of the present invention. Also, FIG. 12 illustrates a process for Bayesian estimation as utilized by the present invention, and FIG. 13 illustrates the use of multichain monitored algorithms useful for developing solutions.